Application for the NBER Digital Economics and AI Tutorial

Ana Paula Milán Hinostroza
The University of New Mexico
Email: apaulamilan@unm.edu
Website: anamilanhinostroza.com

Course of Study

I am a sixth-year Economics PhD student at the University of New Mexico (UNM). I am currently on the academic job market and expect to graduate in Spring 2027. My graduate coursework has focused on microeconomic theory, econometrics, and applied empirical research. During my PhD, I also took a machine learning course that introduced me to different machine learning methods and their applications, although the course was primarily designed for computer science graduate students. I am interested in building on this training and learning how to incorporate AI into my health economics research. I am interested in building on this training and learning how to incorporate AI into my health economics research. As I prepare to transition to the next stage of my academic career, this is a great opportunity for me to learn how to apply these methods and explore new research directions.

Research Interests

My research lies at the intersection of health and public economics, with a focus on programs and interventions aimed at mitigating the negative externalities of risky health behaviors. My current work examines how public policies, particularly alcohol taxation, influence risky health behaviors and related societal outcomes, including crime. A second strand of my research examines how substance use treatment and healthcare delivery affect subsequent healthcare utilization and spending.

Across these projects, I use quasi-experimental methods and large consumer and administrative datasets to answer questions with direct relevance for public policy. I am increasingly interested in how machine learning and AI could complement these methods, particularly given the amount of information available in large healthcare datasets and the importance of understanding risky health behaviors for prevention and treatment.

Dissertation

My dissertation is titled "Public Policy and Risky Health Behaviors." Throughout this work, I think deeply about how alcohol taxation can mitigate risky health behaviors and their negative externalities, and how substance use treatment influences subsequent healthcare utilization.

My first dissertation chapter, which is also my job market paper, examines how alcohol-specific sales taxes affect household alcohol purchases. While an extensive literature examines

how alcohol excise taxes affect consumption, much less is known about the effects of alcohol-specific sales taxes. Unlike volume-based excise taxes, sales taxes are applied as a percentage of retail prices and may generate different behavioral responses because they are thought to be less salient at the point of purchase.

I address this gap by examining Maryland's 2011 alcohol sales tax increase. Using NielsenIQ household scanner data and a synthetic difference-in-differences design, I estimate that the tax reduced monthly ethanol purchases per adult by approximately 5.5% in the full sample and 9.2% among households that purchased alcohol before the tax increase. Purchases decline across beer, wine, and spirits, with the largest absolute reductions among households with high pre-tax alcohol purchases. I find little evidence that households offset these reductions by substituting across alcoholic beverage types. I find small increases in nonalcoholic beer and wine purchases, but no evidence of changes in tobacco purchases.

Using state-level fatal crash data and a two-stage difference-in-differences estimator, my second dissertation chapter finds that larger beer excise tax increases are followed by larger and more precisely estimated reductions in alcohol-impaired fatalities. These findings provide evidence on how alcohol taxation may reduce the social costs of alcohol consumption, particularly through improvements in traffic safety.

In my last chapter, I use New Mexico Medicaid claims data to study residential substance use treatment and compare outcomes for individuals who complete treatment versus those who only detoxify. I find that treatment completion is followed by reductions in total Medicaid spending, driven primarily by lower inpatient spending, alongside an increase in outpatient spending. These findings suggest a shift toward lower-cost, less acute care settings following treatment completion.

Dissertation Committee:

- Brady P. Horn (chair; email: bhorn@unm.edu)
- Sarah Stith (email: ssstith@unm.edu)
- Xiaoxue Li (email: lix@unm.edu)
- Gabriel Sanchez (external member; email: sanchezg@unm.edu)

Interest in AI and Potential Research Applications

I am interested in attending this tutorial because I want to integrate AI into my future research. While I have some exposure to machine learning through my coursework, I have not yet incorporated these methods into my own research projects. I have several ideas for how AI could contribute to my research agenda, although I am still exploring their feasibility.

First, I have an established research pipeline with my advisor using New Mexico Medicaid claims linked to criminal justice data. We plan to study whether residential substance

use treatment affects subsequent criminal justice involvement and how healthcare received during incarceration shapes Medicaid utilization after release.

I am interested in exploring whether AI could help integrate these administrative datasets, particularly when linking records across different systems. More importantly, I want to learn about the potential of AI-generated synthetic data to make these otherwise difficult-to-access datasets more accessible to researchers. These data contain sensitive information about healthcare utilization, substance use treatment, and criminal justice involvement, which makes access difficult. I want to understand whether synthetic datasets could preserve important statistical relationships while reducing privacy risks, allowing more researchers to study substance use and related health outcomes. I am also interested in learning about the limitations of synthetic data, particularly whether they can preserve the relationships needed for causal inference.

My second project will examine how healthcare received during incarceration relates to Medicaid utilization after release, with particular attention to transitions between correctional facilities and community-based care. Beyond creating synthetic datasets for future research, I am interested in exploring whether machine learning could help identify patients who may be at greater risk of gaps in care or adverse health outcomes following release. Understanding which individuals are most vulnerable during these transitions could help inform policies aimed at improving continuity of care.

Although my applications of machine learning have so far been limited to coursework, I want to learn more about how these tools could complement the econometric methods I already use. I am particularly interested in understanding how to evaluate the quality and privacy of synthetic data and how to use machine learning alongside causal inference methods. This tutorial would be a great opportunity to learn from researchers working in this field, better understand which of my ideas are feasible, and identify new research questions that I may not have considered, especially now that I am in the job market. Ultimately, my goal is to use AI to help answer important questions in health economics, while also exploring ways to make sensitive data more accessible to researchers without compromising individuals' privacy.